A study on Privacy-Preserving Federated Learning and enhancement through Transfer Learning

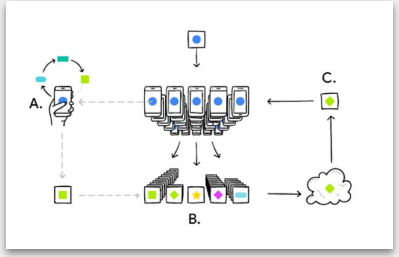
Name: Robert Minea
Student mail: r.minea@student.tudelft.nl
Responsible Professor: Prof.Dr. Kaitai Liang
Supervisor: Rui Wang

1.Background

Neural Networks:
Learning a new task by training a model on a dataset.

Federated Learning:
Decentralized learning done on the datasets of each client. The learning is round based, each client trains their model and then sends the model to the server to be aggregated and returned for the next training round.

Transfer Learning:
The learning on the target network can use the knowledge from a source network, it can either use parts of their trained model or an adapted domain. The techniques for transfer are diverse and not limited to the ones above.



The Federated Learning process from the Google AI blog [1]

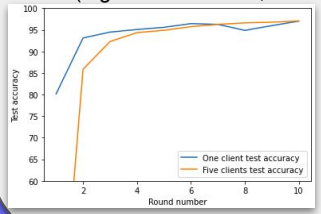
2.Research Questions

- Does Federated Learning fare better than Centralized Learning when it comes to efficiency and accuracy?
- What are the frameworks that allow the usage of transfer learning together with Federated Learning?
- What complexities do the algorithms used have?
- What are the security/privacy techniques used?

3.Research method

Papers presenting different frameworks have been analyzed from a complexity, communication cost and security/privacy point of view.

Attacker types: Honest but Curious, Malicious
Defense Mechanisms: Homomorphic encryption, Differential Privacy, Secure Multi Party Computing, Knowledge Distillation
Attack types: Training attacks (Poisoning), Inference attacks (e.g.Reconstruction, Membership-Inference)



Simulation of a federation has been made in Google Cloud and was compared to a simple Neural Network performance.

4.Results of Framework Study

Algorithm	Security/Privacy Mechanism	Complexity	Communication cost complexity
1.FedAvg	None	Server: O(K • M) Client: O(train)	O(M)
2.FedVote	None	Server: O(K • M + s) Client: O(train + s • (infer + sCKA))	O(M + K)
3.SecureFed Transfer	Homomorphic Encryption/Beaver Triples + Additive secret sharing	Overall: O(R • (Nab • (enc(M1) + enc(M2))))	O(max(enc(M1), enc(M2)) + mask)
4.FedHealth	Homomorphic Encryption	O(R • (K • (enc(train))))	O(enc(M))
5.FedMD	Model Distillation	Server: O(K) Client: O(train)	O(1)
6.FedDistil/FedAug	Knowledge Distillation	Server: O(K • L • logit) Client: O(E • (B • M) +L • logit)	O(logit)
7. Knowledge Federation	SMPC/Homomorphic Encryption/Differential Privacy	-	-

Final summarized table of study results

[1]. Brendan and R. Daniel.Federated Learning: Collaborative Machine Learning with-out Centralized Training Data. 2017.url:https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

TU Delft